AFRL-RH-WP-TR-2012-0067

Elucidation of Small RNAs that Activate Transcription in Bacteria

Michael S. Goodson
Thomas Lamkin
Ryan Kramer

Forecasting Division
Human Signature Branch

John A. Lynch

University of Cincinnati

MARCH 2012
Final Report

AIR FORCE RESEARCH LABORATORY
711TH HUMAN PERFORMANCE WING,
HUMAN EFFECTIVENESS DIRECTORATE,
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the 88th Air Base Wing Public Affairs Office and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-RH-WP-TR-2012-0067 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//signature//

Nancy Kelley-Loughnane, Work Unit Manager
Human Signatures Branch

//signature//

LOUISE A. CARTER, Chief
Forecasting Division
Human Effectiveness Directorate
711th Human Performance Wing
Air Force Research Laboratory

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YY)*
27 03 12

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*
22 April 2008 – 23 January 2012

4. TITLE AND SUBTITLE
Elucidation of Small RNAs that Activate Transcription in Bacteria

5a. CONTRACT NUMBER
FA8650-08-C-6832

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
62202F

6. AUTHOR(S)
Michael S. Goodson[*], John A. Lynch[**], Thomas Lamkin[*], Ryan Kramer[*]

5d. PROJECT NUMBER
2312

5e. TASK NUMBER
A

5f. WORK UNIT NUMBER
2312A218

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
[**]Department of Chemistry
University of Cincinnati
Cincinnati OH 45267

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
*Air Force Materiel Command
Air Force Research Laboratory
711th Human Performance Wing
Human Effectiveness Directorate
Forecasting Division
Human Signatures Branch
Wright-Patterson AFB, OH 45433

10. SPONSORING/MONITORING AGENCY ACRONYM(S)
711 HPW/RHXB

11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)
AFRL-RH-WP-TR-2012-0067

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution A: Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES
88ABW-2012-1307, cleared on 9 Mar 12. Report contains color.

14. ABSTRACT

Small non-coding RNA (sRNA) control of gene expression has been shown to play a prominent role in genetic regulation. While the majority of identified bacterial sRNAs exert their control at the translational level, a few examples of bacterial sRNAs that inhibit transcription have also been identified. Using an engineered combinatorial RNA library, we have elucidated bacterial sRNAs that activate transcription of a target gene in *E. coli* to varying degrees. Mutation of the strongest activator modified its activation potential. Our results suggest that transcriptional activation of our target gene results from recruitment of the bacterial RNA polymerase complex to the promoter region. These data, coupled with the malleability of RNA, provide a context to define synthetic control of genes in bacteria at the transcriptional level.

15. SUBJECT TERMS
Small non-coding RNAs, transcriptional activation, bacteria, signal amplification, synthetic control

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT:**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** (Monitor)
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Nancy Kelley-Loughnane
Unclassified	Unclassified	Unclassified	SAR	43	**19b. TELEPHONE NUMBER** *(Include Area Code)* N/A

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18

THIS PAGE IS INTENTIONALLY LEFT BLANK

TABLE OF CONENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENT

We thank Dr. K. Greis at the University of Cincinnati Cancer and Cell Biology Proteomics Core Facility and Dr. N. Kelley-Loughnane, M. Davidson, A. Stapleton, G. Sudberry, and J. Wright at 711th Human Performance Wing for their assistance. This work is funded by the Bio-X STT, Air Force Research Laboratory.

1.0 INTRODUCTION

The central dogma of biology posits RNA as a conduit for genetic information to flow from DNA to proteins. However, the role of small non-coding RNA (sRNA) has been found to include a number of diverse biological functions, including the regulation of gene expression (Storz, 2002; Storz et al., 2005). Indeed, sRNAs are hypothesized to be essential in the genomic programming of complex organisms (Mattick, 2004). The majority of identified bacterial sRNAs that regulate gene expression do so by either base pairing with other RNAs or binding and regulating protein activity (Storz et al., 2005; Waters and Storz, 2009). Most of these interactions disrupt or modulate RNA translation. However, there is a growing class of bacterial sRNAs that are also known to inhibit transcriptional processes. These sRNAs act by binding to and inhibiting RNA polymerase (Wassarman and Storz, 2000), and by base pairing to mRNA to form a secondary structure that leads to transcriptional termination (Novick et al., 1989, Storz et al., 2006). Activation of transcription by sRNA has been described in eukarytoes (Buskirk et al., 2003; Saha et al., 2003, Wang et al., 2010), but no such activity has been implicated in bacterial gene regulation. Despite this, a number of sRNAs from a fragmented *E. coli* genome library have been found to bind to RNA polymerase with high affinity (Windbichler et al., 2008). Thus, RNA-based regulatory elements may exist within the transcriptional processes of prokaryotes.

Synthetic biology aims to exploit cellular processes by genetically engineering key nodes in transcriptional and translational processes. The relative ease of engineering, screening, and modeling of RNA compared to that of proteins (Buskirk et al., 2003; Culler et al., 2010; Liu and Arkin, 2010) uniquely identifies these biomolecules as highly engineerable elements for exploitation in synthetic systems. The most prominent methods for synthetic sRNA-based regulation of bacterial gene expression involve translational modulation via ‘riboswitches’

(Isaacs et al., 2004; Harbaugh et al., 2008; Topp and Gallivan, 2010), binding and inhibition of a specific bacterial transcription repressor (Hunsicker et al., 2009), or attenuating transcription through structural changes in RNA effected by antisense RNA (Lucks et al., 2011). Up-regulating transcription should increase reporter protein levels by orders of magnitude compared to translational control elements alone, where the amount of protein transcribed is dependent upon the number of transcripts available. This is particularly advantageous in engineered bacterial systems designed as cell-based reporters for environmental monitoring. The ability to exert both transcriptional and translational control within an engineered circuit also enables a number of different capabilities. For instance, redundancy in detection at both the transcriptional node and translational node could decrease false positives. Engineering different ligand detection schemes into each node also allows for more complex logic in biologically-based detection and reporter systems.

Here we describe the first identification and characterization of bacterial sRNA transcriptional activators using an *in vivo* selection technique for prokaryotes, and consequently add a new component to the tool kit of synthetic biology. A similar approach identified RNA-based transcriptional activators in yeast (Buskirk et al., 2003) although the underlying mechanism was not elucidated. We show that the bacterial RNA transcriptional activators characterized here associate with the RNA polymerase complex of proteins, and we propose that the RNA initiates transcription through recruitment of the RNA polymerase holoenzyme.

2.0 RESULTS AND DISCUSSION

2.1. Identification of potential RNA transcriptional activators

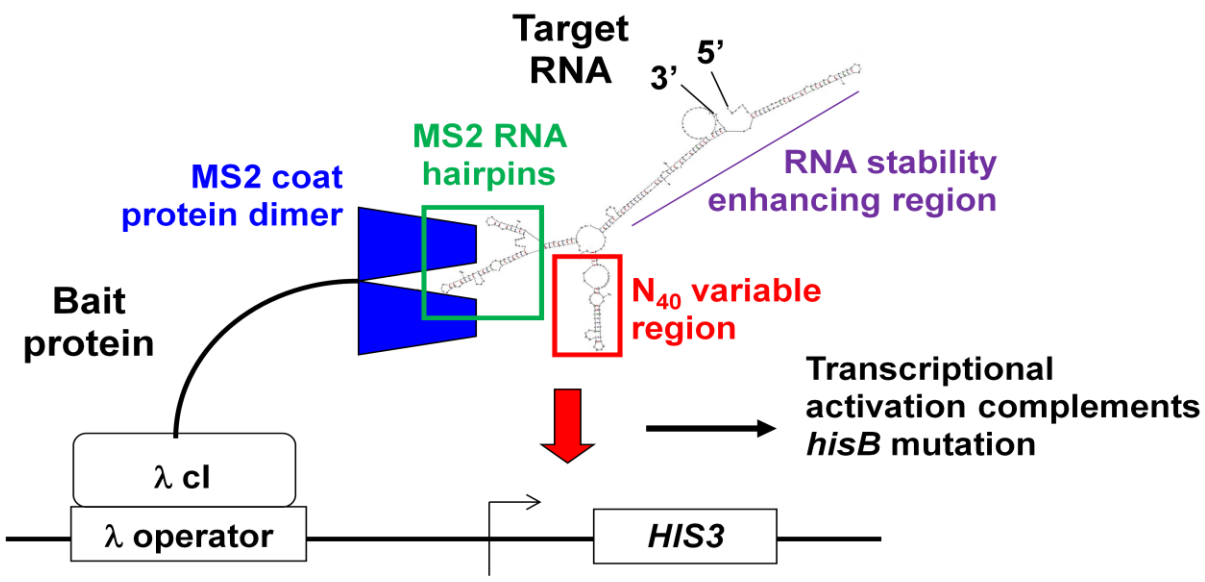
In this study, we modified a commercially available bacterial two-hybrid system (BacterioMatch II two-hybrid system vector kit; Stratagene, La Jolla, CA) for the rapid selection of a RNA-based transcriptional activation sequence. Instead of monitoring protein-protein interaction via the assay, we exploited the high affinity interaction between the MS2 coat protein and the MS2 RNA hairpin. The bait plasmid (pBT) was modified by inserting the coding sequence for a MS2 coat protein dimer (Genescript, Piscataway NJ) into the 3'-end of a λcI coding sequence already present within the pBT plasmid. The resultant plasmid was termed pBT-MS2. The functional fusion protein λcI-MS2 binds to the lambda operator located upstream of a *HIS3* reporter gene via the DNA binding domain of λcI (Fig. 1A). The commercial target vector (pTRG) was modified by replacing the coding region for the RNA polymerase gene with a non-coding region that contained three distinct elements (See Fig. 1A). First, a RNA with a known stable secondary structure was included at the 5' and 3' ends to minimize transcript degradation (Buskirk et al., 2003). Second, two MS2 RNA hairpins were included that localize the transcript to the promoter via interaction with the λcI-MS2 fusion protein. Third, a variable library consisting of 40 randomly synthesized nucleotides was cloned upstream of the MS2 binding sequence. This element was screened for transcriptional activation at the reporter site. The resultant plasmid, containing all three components, was termed pTRG-var.

Both pBT-MS2 and pTRG-var were cotransformed into *hisB* knock-out *E. coli* containing an F' episome which included the reporter gene cassette. Growth of these bacteria on histidine-free minimal media supplemented with 3-amino-1,2,4-triazole (3-AT) was used as an

indicator of *HIS3* transcriptional activation. 3-AT acts as a competitive inhibitor of the low levels of *HIS3* gene product produced by the reporter gene cassette even in the absence of transcriptional activation. As a result the untransformed reporter strain was unable to grow on media lacking histidine in the presence of 3-AT. Similar approaches have been successfully utilized in previous yeast two-hybrid (Buskirk et al., 2003) and three-hybrid studies (Bernstein et al., 2002).

Colonies were observed from co-transformations of pBT-MS2 and pTRG-var plated onto 2 mM 3-AT selective screen plates. Compared to colonies grown on non-selective media, results indicated that ~0.1% of clones could grow on selective screen plates, a proportion similar to that found in a yeast model system (Buskirk et al., 2003). A total of 22 positive co-transformations were identified, and all 22 grew on 3-AT selective screen plates after plasmid isolation and re-co-transformation. No colonies were observed from co-transformations of the unmodified pBT plasmid and a positive pTRG-var. This suggested that localization of the RNA transcriptional activator to the reporter gene is necessary for transcriptional activation. Similarly no colonies were observed from co-transformations of pBT-MS2 and pTRG that were plated onto selective screen plates. Alignment of the variable regions of the RNA constructs from these co-transformations exhibited large sequence variation without any obvious conserved domains, as evidenced by the long branch lengths and low bootstrap values in Fig. 1B. Secondary structure analysis (mfold; Zuker, 2003) indicated that the variable regions of all the RNA transcriptional activators identified produced a structure in the RNA molecule that could be grouped into those producing a single, double, or paired stem-loop (Fig. S1). These data suggest that transcriptional control by sRNA is prevalent in bacteria, and that multiple biological interactions may drive transcriptional activation at the reporter site.

A



B

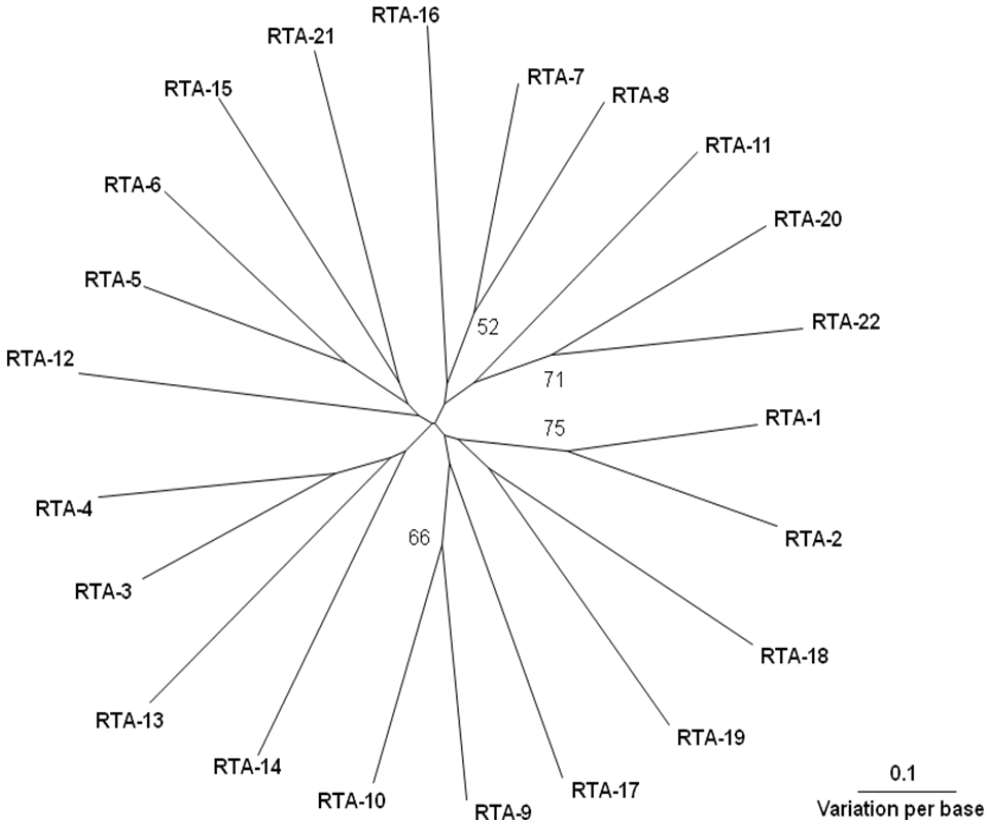


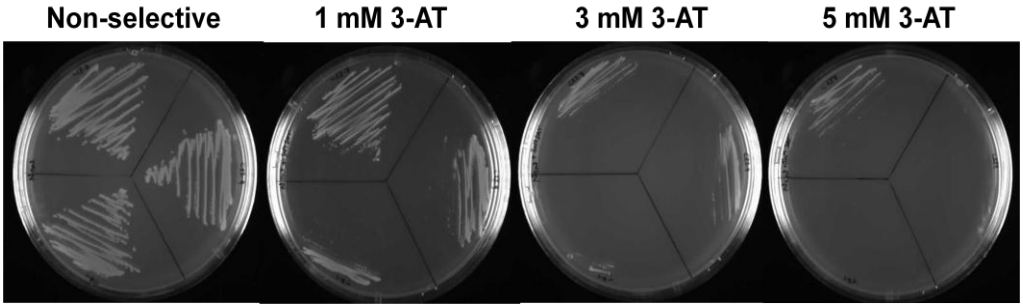
Figure 1. Screening for RNA transcriptional activation

A) The bait plasmid carries λcI genetically fused to a recombinant MS2 coat protein sequence which results in the translation of the fusion protein (top). The target plasmid contains three elements, an RPR region for transcript stabilitity, a MS2 sequence which localizes the transcript to the λcI-MS2 protein, and a N40 variable region. The sequence contains no translational signals and is maintained intracellularly as an RNA transcript (middle). Only those RNA library members that activate transcription of HIS3 allow survival on selective screen media (bottom). Modified from Buskirk et al., 2003. B) Unrooted phylogenetic tree produced by the neighbor-joining method, using Kimura two-parameter distances between the variable regions of RNA molecules that could activate transcription of HIS3. Gapped positions were included in the analysis. Percentage bootstrap values >50 are indicated. See also Fig. S1.

2.2 Selection of strong transcriptional activators

There was variation in the ability of the positive co-transformations to grow on plates containing increasing concentrations of 3-AT (Fig. 2A). This suggested that there may be variation in the overall strength of transcriptional activation. To quantitatively address these variations, an initial selection procedure involving quantitative real-time PCR (QPCR) analysis and growth kinetics analysis (Fig. S2) was employed. Comparison of these methods identified statistically similar patterns of RNA transcriptional activation strength ($\rho = 0.689$, $P<0.0005$). RNA transcriptional activators were categorized as 'Strong', 'Medium', or 'Weak' based on comparing their ranks in the growth kinetics and the initial QPCR experiments. There was no obvious association between strength of activation and predicted secondary structure. Representatives of the Strong, Medium, and Weak RNA transcriptional activators were used in two additional biological replicates of the QPCR analysis. For all QPCR experiments, a low variation in Ct value was observed within each triplicated sample. 'Strong' RNA transcriptional activators had consistently higher fold differences than 'Medium' and 'Weak' RNA transcriptional activators. There is a significant difference in activation strength among the Strong RNA transcriptional activators (ANOVA, $P<0.0005$) (Fig. 2B).

A



B

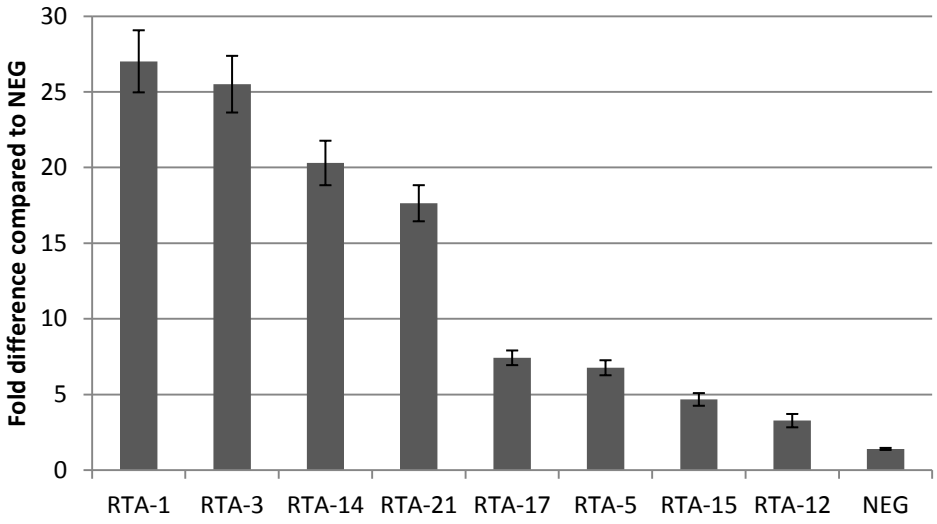


Figure 2. Assessment of the strength of transcriptional activation

A) Variation in the strength of transcriptional activation as determined by the ability to grow on increasing concentrations of 3-AT. Strains were streaked onto plates and incubated at 37°C for 24 h. Upper segment: RTA-3; right hand segment: RTA-5; lower segment: RTA-12. B) Average fold difference of HIS3 expression normalized to RNA transcriptional activator expression of STRONG, MEDIUM, and WEAK transcriptional activators compared to NEG. Fold difference was compared using the delta delta Ct method. Each individual Ct value of three replicates of three biological replicates per transcriptional activator were compared to the mean NEG Ct to calculate each delta delta Ct value (n=81). Data are represented as mean +/- SEM.

The two strongest RNA transcriptional activators, RTA-1 and RTA-3 are significantly different from RTA-21 (Tukey's pairwise comparison, $P<0.01$). There is no significant difference between RTA-1 and RTA-3 (t test, $P>0.2$). As a further selection procedure, the growth of each of the four strongest RNA transcriptional activators was assessed on 5 mM and 3

mM 3-AT selective screen plates (Table 1). These concentrations of 3-AT were more stringent than for the initial screening of pBT-MS2 and pTRG-var cotransformants. Selectivity was confirmed by the growth of the positive control co-transformed *E. coli* (pBT-LGF2 pTRG-Gal11^{P}) on all of the selection plates after 18 h (Table 1), while the negative control co-transformed *E. coli* only grew on the non-selective screen plates. Three of the four strong RNA transcriptional activators, RTA-3, RTA-14, and RTA-21, grew on the most stringent (5 mM 3-AT) selective screening plate after 24 h. However, of these three, RTA-21 grew only slowly, as evidenced by the formation of small colonies. These three strong RNA transcriptional activators exhibited colony growth on the 3 mM selective screen after 24 h incubation, although the colonies from RTA-21 were small. RTA-1 exhibited colony growth on the 3 mM selection plates after 32 h incubation. All samples grew on the non-selective screen after 18 h incubation.

Table 1. Assessment of growth of the strongest RNA transcriptional activators on plates of differing selection stringency

Sample	5 mM 3-AT			3 mM 3-AT			No 3-AT		
	18 h	24 h	32 h	18 h	24 h	32 h	18 h	24 h	32 h
RTA-3	x	✓✓	✓✓	x	✓✓	✓✓	✓✓	✓✓	✓✓
RTA-1	x	x	x	x	x	✓	✓✓	✓✓	✓✓
RTA-14	x	✓✓	✓✓	x	✓✓	✓✓	✓✓	✓✓	✓✓
RTA-21	x	✓	✓	x	✓	✓	✓✓	✓✓	✓✓
Positive	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓
Negative	x	x	x	x	x	x	✓✓	✓✓	✓✓

x: no colonies
✓: small colonies
✓✓: large colonies

In order to determine the strongest RNA transcriptional activator, their order in the growth kinetics analysis, the compiled real-time PCR data, and results from the selective screen assay were converted to a rank. For the latter, rank was attributed using qualitative growth rates

on the 5 mM 3-AT selective screen. Using this criteria, the strongest RNA transcriptional activator was determined to be RTA-3 (Table 2). This RNA had no similarity to sequences present in the *E. coli* genome (BLAST; Altschul et al., 1990). RTA-3 was subsequently used in mutational studies aimed at determining sequence relevancy and analyses to determine the possible interacting proteins that drive transcriptional activation.

Table 2. Selection of the strongest RNA transcriptional activator

Sample	Rank			Total
	Growth rate constant, k (Fig. S2)	Compiled QPCR (Fig. 2C)	Selective screen assay (Table 1)	
RTA-3	1	2	1	4
RTA-1	2	1	4	7
RTA-14	4	3	1	8
RTA-21	3	4	3	10

2.3 Identification of interacting proteins

To investigate which proteins potentially interact with the RNA transcriptional activator, an electrophoretic mobility shift and protein sequencing assay was employed. A biotinylated, RNA probe specific to RTA-3 was used to qualitatively determine if the RNA transcriptional activator interacted with native *E. coli* proteins. Mobility shifts of the biotinylated RNA probe were seen in native gel electrophoresis of co-transformed RTA-3 bacterial cell extract (Fig. S3). Sequencing of the regions of gel showing a probe size shift returned a total of 37, 61, and 89 proteins with significant ($P<0.005$) matches to the NCBI database, for regions 1, 2, and 3, respectively (Table S1). Three members of the RNA polymerase holoenzyme were abundant in regions 1 and 2. RNA polymerase beta (rpoB) was the second and third most abundant protein in regions 1 and 2, respectively. RNA polymerase sigma 70 (rpoD) was the 12th most abundant

protein in region 1, and RNA polymerase alpha (rpoA) was the 17th and 15th most abundant protein in regions 1 and 2, respectively. The rpoB catalyzes the synthesis of RNA, while rpoA assembles the enzyme and binds regulatory factors. The rpoD is required to enable specific binding of RNA polymerase to the promoters of most genes required in growing cells (Ptashne and Gann, 1997). The lambda repressor protein, λcI was identified in region 3, suggesting that the interaction between bait and target complex remained intact under native electrophoresis. These analyses suggested that our identified transcriptional activator RTA-3 may recruit the polymerase complex to the reporter gene promoter region.

To test this conclusion, we investigated whether the RNA transcriptional activator could be co-immunoprecipitated with RNA polymerase subunits using antibodies to rpoB and rpoD. To account for non-specific RNA binding, immunoprecipition experiments included cell extracts from bacteria transformed with pBT-MS2 and a target plasmid that was unable to activate transcription, pTRG-control. This target plasmid was identical to pTRG-var except it contained a ten nucleotide insert instead of the 40 nucleotide variable region. As expected, each subunit was immunoprecipitated from cell lysates of bacteria co-transformed with pBT-MS2 and pTRG-varRTA3, as well as bacteria transformed with pBT-MS2 and pTRG-control, and transformed with pBT-MS2 alone. Reverse transcription PCR (RT-PCR) was performed using primers specific to the transcriptional activator to identify if RTA-3 was co-precipitated with rpoB and rpoD. The RNA transcriptional activator RTA-3 was detected in the rpoB and rpoD precipitated fraction from bacteria cotransformed with the bait plasmid and pTRG-RTA-3 (Fig. 3). No RT-PCR product was seen in the rpoB and rpoD precipitated fraction from bacteria co-transformed with the bait and pTRG-control plasmid or the bait plasmid alone. PCR analysis determined that there was no bait or target plasmid DNA contamination of the eluate which may have given rise

to a false positive (Fig. 3). Taken together, these results indicate that the RTA-3 RNA transcript was able to be co-immunoprecipitated with the RNA polymerase subunits, rpoB and rpoD. Direct or indirect recruitment of RNA polymerase to the promoter region of the reporter by the RNA transcriptional activator would have the direct effect of stimulating or enhancing transcription from this site.

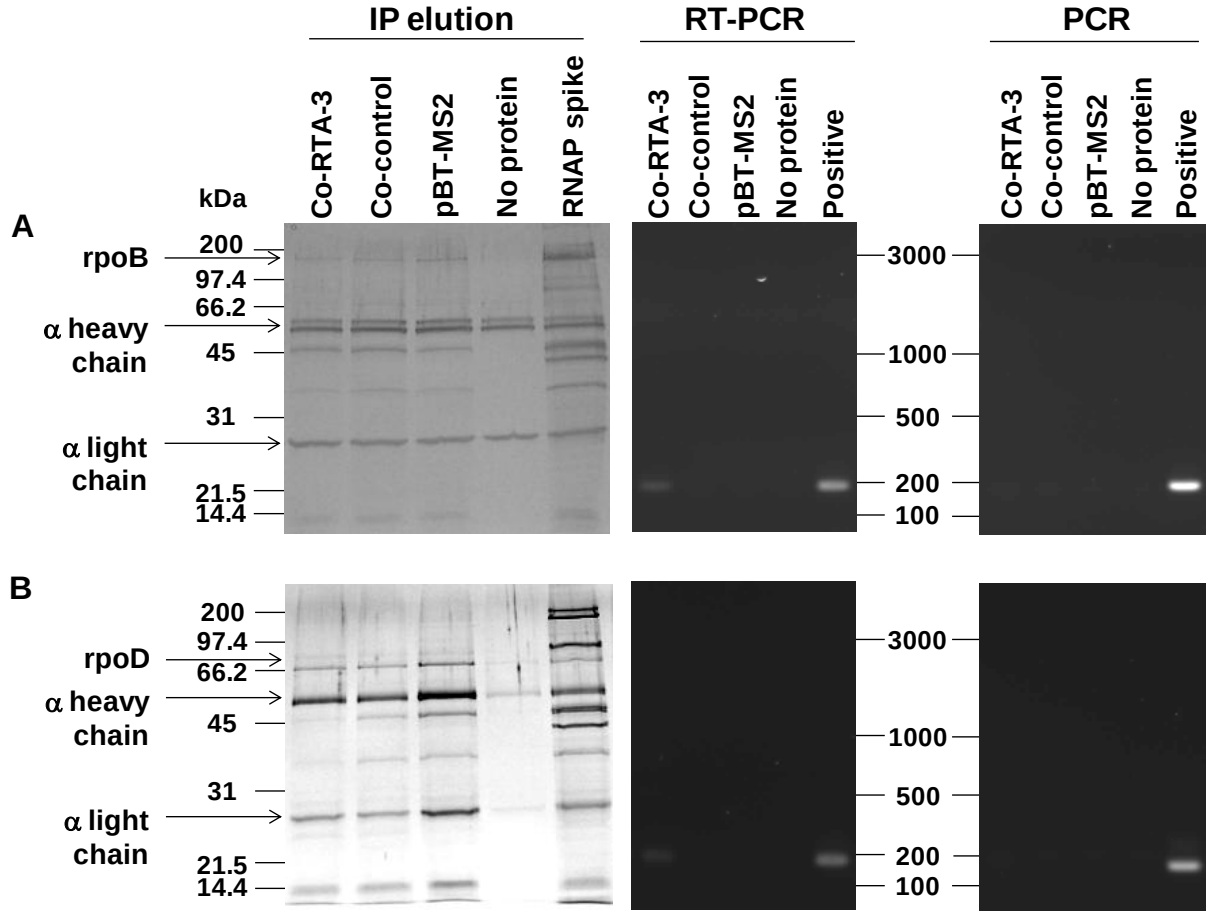


Figure 3. Immunoprecipitation of RTA-3 with rpoB and rpoD

Reverse transcription PCR and PCR of the eluate from the rpoB immunoprecipitation (A) and the rpoD immunoprecipitation (B) using RNA transcriptional activator-specific primers. Amplicon sizes of 193 bp and 163 bp are expected for RTA-3 and control, respectively . No DNA contamination of the eluate by the target plasmid was detected. No Protein: control immunoprecipitaton experiment performed without the addition of cell lysate; RNA spike: lysate from co-RTA-3 culture spiked with RNA polymerase holoenzyme; Pos RTA-3: reactions performed with purified pTRG-RTA-3 as the template; Pos control: reactions performed with purified pTRG-control as the template. Gels are representative of three independent experiments.

2.4 Mutagenesis of the strongest RNA transcriptional activator

The sequence variability exhibited in the identification of potential RNA transcriptional activators in our initial screen and the differences in their strength of activation suggests that modification of a specific RNA transcriptional activator will affect its activation capability. To test this hypothesis we investigated the effect of randomly mutating the 40 nucleotide insert region of the strongest transcriptional activator, RTA-3, to determine if the strength of transcriptional activation could be changed, and to elucidate which parts of this region were essential for activation. The resultant clones were subjected to the selection procedures described above. An average of 4.9 (SE = 0.5) mutations per clone was achieved. Twenty six mutagenized variants of RTA-3 were able to grow on 3-AT selective screen media. The consensus sequence of the aligned mutations (Fig. 4A) was identical to the variable region of RTA-3, suggesting that RTA-3 is already relatively well optimized in its ability to activate transcription. However, there was only one conserved residue in all the mutant clones (A30) which closed a loop near the end of the single stem predicted to form in RTA-3. Thirteen additional sites showed one or more substitutions of the same nucleotide, including a region of six consecutive bases (19-24). This region corresponds to the terminal loop of the single stem formed by RTA-3 (Fig. S1). To assess the strength of activation, QPCR was performed as described above except that fold difference was expressed relative to the level of transcriptional activation exhibited by RTA-3 (Fig. 4B). Six mutations exhibited a greater than 3-fold increase in transcriptional activation compared with RTA-3, with the strongest activator, 3-M6, exhibiting an 8-fold increase. This is comparable to a similar study performed in yeast (Buskirk et al., 2003), although we found less sequence conservation between mutants than in that study. Two mutant clones 3-M6 and 3-M17, are significantly stronger transcriptional activators than the

others (ANOVA; $P<0.0001$), although they are not significantly different from each other (t test; $P>0.1$). Interestingly, two of the three strongest transcriptional activators, 3-M17 and 3-M3, differed from RTA-3 by only a single base. There was no obvious effect of secondary structure on activation strength, although mfold predictions of the mutated activators showed that the two mutant clones exhibiting the strongest transcriptional activation, 3-M6 and 3-M17, had a similar triple stem-loop structure despite their sequence differences (Fig. S4). The only other mutant clone predicted to have a similar triple stem-loop structure, 3-M21, exhibited weaker transcriptional activation than the original RTA-3 clone. The semi-conserved consecutive nucleotides 19-24 of the 40 nucleotide insert formed the terminal loop of the middle stem in all three clones. However, the length of the predicted stems were shorter in 3-M21 than those of 3-M6 and 3-M17. The mutant clones exhibiting the weakest transcriptional activation compared to the original RTA-3 clone, 3-M13, 3-M20, and 3-M16, had 8, 3, and 3 substitutions, respectively. Similarly, there was no obvious difference in their predicted secondary structure compared to clones that had a higher transcriptional activation capability than RTA-3 (Fig. S4). These data suggest that both sequence *and* structure are important determinants of activation strength. This, coupled with the requirement for localization of the RNA transcriptional activators upstream of the reporter gene in our system, supports the recruitment model of transcriptional activation (Ptashne and Gann, 1997) in bacteria.

A

```
                    1    1    2    2    3    3    4
               1    0    5    0    5    0    5    0
                -   -   -  -   - ------  -  *   -
3-M1           CGTCTCAGTTCTGATAATACGAGGGTGGCAGTTATCGC--
3-M2           CGCCTCGGTCTCGGTAATACGGGAGTGGCAGCTCTCGCCT
3-M3           CATCTCAGCCCCGGTAATACGAGAGTGGTAGTTATCGCCC
3-M4           CATCTCAGCCCCCGCAATACGAGAGTGGCAGTTATCGCCC
3-M5           CACCTCAGCCCCGGTAATACGAGAGTGGTAGTCATCGCCA
3-M6           CACCTCAGCCCCGGTAGTACGAAAATGGTAGTTATCGCCA
3-M7           CACCTCAGCCCCGGTAATACGGGAGTGGTAGTTATCACCA
3-M8           CATCTCAGCCCCGGTCAATCTAGAGTTGTAGTTATAGCCA
3-M9           TATCTCAGCCCCGGTGGTACGAGAGTGGTAGTTATAACCA
3-M10          CATTACAGCCCTGGTAATACGAGACTGATAGTTATAGCCA
3-M11          CATTACAGCCCTGGTAATACGAGACTGATAGTTATCGCCA
3-M12          CATCTCTGCCCGGGAAACACGAGGGTGGTAGCAATCCACA
3-M13          GATCAGTGCCCAGGTAATACGAGAGGGGTAGTTATCCACA
3-M14          CATCCCAGCCCCGGTAATACGAGAGTGGCAGTTATCGCCA
3-M15          CATCTCAGGTGCGCTAATACGAGAGTGCAAGTTATCGCCA
3-M16          CATCTCAGGCCCGGTAATACGAGAGTGGTAGGGATCGCCA
RTA-3          CATCTCAGCCCCGGTAATACGAGAGTGGTAGTTATCGCCA
3-M17          CATCTCAGCCCCGGTAATACGAGAGTGGTAATTATCGCCA
3-M18          CATATCAGCCCCGGCAATACGAGAGTGGTAGTTATCCCGA
3-M19          CAACTCTACCCCGGTAATACGAGAGGGGTAGTTATCTCGA
3-M20          CACCTGACCCCCGGTAATACGAGAGTGGTAGTTATCGCCA
3-M21          CAATTCACCCCCGGTAATACGAGAGTGGTACTTACCGCCA
3-M22          CATCTCACCCCCGGCAAAACGAGAGTGGTAGATAACGCCA
3-M23          CATATCAGCCCCGGGAATACGAGAGTGGTAGTCATCGCCA
3-M24          CGTCTCAGCCCCGGTAATACGAGAGGGGTAGTTACCGCCA
3-M25          CATCTCCTCCCCGGTAATATGAGAGCGGCAGTTATGGCCA
3-M26          CGTCTCGACCCCGGTAGTACGGGAGTGGTAGTTATGATGC

Concensus      CATCTCAGCCCCGGTAATACGAGAGTGGTAGTTATCGCCA
```

B

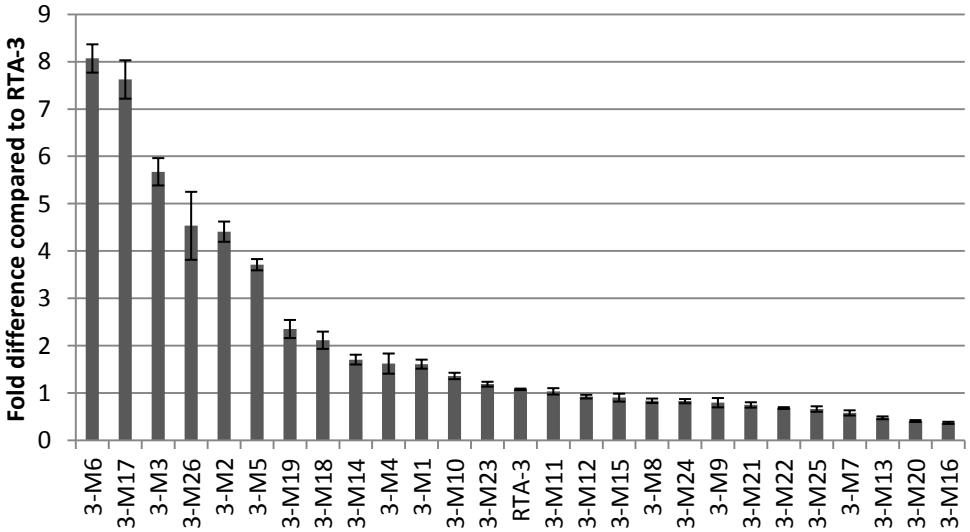


Figure 4. Evolution of stronger activators

A) Alignment of mutated RTA-3 sequences that grow on 3-AT selective screen media. Stars indicate conserved nucleotides. Dashes indicate columns that have one or more substitutions of the same nucleotide. B) Average fold difference of HIS3 expression normalized to RNA transcriptional activator expression compared to RTA-3. Fold difference was compared using the delta delta Ct method. Each individual Ct value of three replicates per mutant were compared to the mean RTA-3 Ct to calculate each delta delta Ct value (n=81). Data are represented as mean +/- SEM. See also Fig. S4.

RNA molecules share many chemical features found in protein transcriptional activators (Buskirk et al., 2003) and these features may be sufficient to allow non-natural RNA sequences to interact with RNA polymerases to mediate transcription. Our experiments show that the RTA-3 RNA transcriptional activator does associate with the bacterial RNA polymerase complex and it is likely that recruitment of the RNA polymerase by the activator is sufficient to activate transcription of the adjacent reporter gene, in line with the recruitment model of transcriptional activation. Protein based, natural bacterial transcriptional activators have been shown to bind to different parts of the RNA polymerase holoenzyme to activate transcription (Bushman et al., 1989, Busby and Ebright, 1994, Ptashne and Gann, 1997). While we have not been able to identify which part of the bacterial RNA complex our strongest RNA transcription activator associates with, the recruitment model suggests that a specific interaction or recognition site is not a requirement of activation. It is not inconceivable to suggest that the RNA transcriptional activator may interact with multiple parts of the RNA polymerase complex. Indeed, the variation in strength of activation may be associated with the number of contact sites between the activator and polymerase (Ptashne and Gann 1997), with increased contact sites increasing the activation strength, potentially through higher binding leading to increased 'ON' time for the polymerase in this stochastic system. The variation in strength of activation of the mutated RTA-3 activator suggests that even small changes in sequence and subtle changes in

secondary structure can affect the activation strength, possibly by affecting interactions with the RNA polymerase complex.

By selecting for sequences that activate transcription from a random RNA library, and then mutagenizing the strongest activator to change its activation strength, we have increased the repertoire of building blocks for use in synthetic biology. These RNA molecules perform 'protein-like' functions, without the difficulty and *in vivo* immunogenic limitations of proteins (Wang et al., 2010). In the same way that the modular components of eukaryotic transcription factors can aid in the synthesis of genetic switches (Ptashne and Gann, 1997), coupling two RNA building blocks consisting of an aptamer specific to a molecule of interest and a RNA transcriptional activator would allow controlled regulation of transcription. RNA-based ligand-dependent transcriptional activation has been shown to be achievable in yeast (Buskirk et al., 2004), and our findings suggest that a sRNA with a similar function can be engineered in bacteria.

3.0 METHODS

All chemicals were purchased from Sigma-Aldrich (St. Louis, MO) and all restriction enzymes were purchased from NEB (Ipswich, MA), unless otherwise stated. All media was prepared as described in the BacterioMatch II 2-hybrid system vector kit protocol (Stratagene, La Jolla, CA).

3.1 Synthesis of the target and bait vectors

The commercial target vector, pTRG, from the BacterioMatch II two-hybrid system vector kit (Stratagene , La Jolla, CA) was modified by adding a MfeI restriction site upstream of the pTRG promoters lpp and lac-UV5 (pTRG-MfeI) using the QuikChange kit (QIAgen, Valencia, CA).

This allowed a synthesized stretch of DNA consisting of an MfeI restriction site, the promoters lpp and lac-UV5, a RNase P leader, BamHI and NotI restriction sites, MS2 hairpins, a RNase P terminator , and a XhoI restriction site (Genscript, Piscataway, NJ) to be inserted in place of the pTRG RNA polymerase alpha gene. Briefly, the synthesized DNA and pTRG-MfeI were double digested using MfeI and XhoI, and digested products of the expected size were gel purified using QIAquick spin minicolumns (QIAgen, Valencia, CA) and ligated using T4 DNA ligase (3 U/ml; Promega, Madison, WI). Ligation products, uncut pTRG (positive control) and gel-excised, double digested pTRG alone (negative control) were transformed into *E. coli* XL1 Blue MRF' Kan competent cells by heat shock following the protocol described by the BacterioMatch II manual (Stratagene, La Jolla, CA). Transformation mixtures were plated onto LB-tetracycline selection plates. Insertion of the synthesized DNA was confirmed by PCR, restriction digest, and sequencing of plasmids purified from transformed colonies. Sequencing was performed by Genewiz (South Plainfield, NJ). Sequences were aligned using ClustalX (Thompson et al., 1997). This modified target vector is henceforth referred to as 'pTRG-MS2'.

A blunt-ended double-stranded variable library was synthesized from a 70 nucleotide primer (Integrated DNA Technologies, Coralville, IA), containing a 40 nucleotide variable sequence between BamHI and NotI restriction sites, using a specific reverse primer and the Klenow fragment of *E. coli* DNA polymerase I (NEB, Ipswich, MA). These reactions and the modified commercial target vector, pTRG-MS2, were double digested using BamHI and NotI. Products were gel purified using the QIAEX II system (QIAgen, Valencia, CA) and ligated using T4 DNA ligase (Promega, Madison, WI). Ligations were transformed in duplicate into XL1-Blue MRF' Kan strain *E. coli* by the heat shock technique. The resultant transformation culture was plated onto LB-tetracycline selection plates. Transformation using an uncut pTRG plasmid

and a ligation reaction containing no plasmid were performed as positive and negative controls, respectively. Insertion of the variable library was confirmed as described above. No similarity was exhibited by 25 clones sequenced to assess library diversity. This modified target vector is henceforth referred to as 'pTRG-var'.

The commercial bait vector, pBT, was modified by inserting a synthesized stretch of DNA encoding a MS2 coat protein dimer, terminating with NotI and XhoI restriction sites (Genescript, Piscataway, NJ) into the multiple cloning site. Digested products of the expected size were gel purified using QIAquick spin minicolumns (QIAgen, Valencia, CA) and ligated using T4 DNA ligase (3 U/ml; Promega, Madison, WI). Ligation products, uncut pBT (positive control), and gel-excised double digested pBT alone (negative control) were transformed into *E. coli* XL1 Blue MRF' Kan competent cells by the heat shock technique. Transformation mixtures were plated onto LB-chloramphenicol selection plates. Insertion of the variable library was confirmed by as described above. This modified bait vector is henceforth referred to as 'pBT-MS2'.

3.2 Expression of pTRG-var and pBT-MS2

Expression of the inserted DNA in the pTRG-var plasmid can be induced by IPTG. RNA expression was confirmed by growing *E. coli* XL1-Blue MRF' cells transformed with pTRG-var in LB tetracycline + IPTG at 30^{o}C, extracting RNA, and detecting presence of the transcript by reverse transcription-PCR (RT-PCR). RNA was extracted using the MasterPure Complete DNA and RNA Purification Kit (Epicentre Biotechnologies, Madison, WI), following the manufacturer's protocol for cell samples, except that double the suggested amount of DNase I was used. RT-PCR was carried out using primers (RTRNA-F, GGCTAGAACTAGTGGATCC, Tm 51.5 ^{o}C and RTRNA-R, TTGGATATGGGGGAATTCC, Tm 51.7^{o}C) that anneal to regions

within the insert using the Access RT-PCR System (Promega, Madison, WI). Reactions were carried out using 20-50 ng total RNA and an amplification annealing temperature of 51^{o}C. Results were analyzed by gel electrophoresis. No DNA contamination was detected using reactions without reverse transcriptase.

IPTG induction of pBT-MS2 should result in expression of a protein encompassing both the λcI and the MS2 coat protein dimer. Protein expression was confirmed by growing *E. coli* XL1-Blue MRF' cells transformed with pBT+MS2 in M9+ His-dropout broth + chloramphenicol and IPTG, extracting protein, and performing a western blot using an antibody to the λcI protein (Stratagene, La Jolla, CA). Protein was extracted using the BugBuster protein extraction reagent (EMD Biosciences, San Diego, CA), following the manufacturer's protocol. Soluble protein quantities were assessed using the BCA protein assay (Thermo Scientific) using a Nanodrop spectrophotometer. SDS-PAGE and western blots were performed using the NuPAGE Novex Bis-Tris gels, the XCell Surelock Mini Cell electrophoresis apparatus and the XCell II Blot module (Invitrogen, Carlsbad, CA). Cross-reactive bands were detected using the ECL plus western blotting detection system (Amersham, GE Healthcare, Piscataway, NJ), following the manufacturers protocol, and photographic film and developer (Kodak, Rochester, NY).

3.3 Identification of potential RNA transcriptional activators

Cotransformation was performed by following the BacterioMatch II two-hybrid system vector kit protocol (Stratagene, La Jolla, CA). Briefly, Bacteriomatch II validation reporter competent cells were transformed by heat shock with the bait and target vector, and plated onto non-selective screening plates and 2 mM amino-1,2,4-triazole (3-AT) selective screening plates. Cotransformation was performed using 50 ng pTRG-var and pBT-MS2, and also 50 ng positive control target (pTRG_Gal11^{P}) and bait (pBT_LGF2) vectors supplied with the BacterioMatch II

kit. Additionally transformations using 50 ng pBT-MS2 alone, and a no vector control were also performed. The number of colonies on each plate was assessed after incubation at 37^{o}C for 24 h. An additional 24 h incubation was performed in the dark at room temperature to allow growth of cells harboring weak interactors and/or expressing proteins that may be toxic to the cells. As expected, the proteins expressed by the positive control bait and target plasmids interact strongly and therefore a similar number of colonies (10^4-10^5 cfu per plate) were visible on the non-selective and selective screening plates. The pTRG-var pBT-MS2 cotransformation had a transformation efficiency of ~10^6 cells per mg, based on growth on nonselective screening plates. The 'no vector' control and the bait vector alone produced no colonies on the nonselective or selective screening plates. Colonies were screened on 54 selective screening plates. The average number of co-transformed colonies counted on nonselective screening plates was $7x10^3$, suggesting that potentially $4x10^5$ colonies were finally screened. Colonies that grew on selective screen plates were further verified by performing re-co-transformation of pBT-MS2 and the pTRG-var vector isolated from the original potential positive clone. Plates were incubated at 37^{o}C for 24 h followed by incubation at room temperature in the dark for 24 h. Verified positives were sequenced in both directions and analyzed as described above. A bootstrapped (1,000 iterations) neighbor-joining tree of the positive sequences was constructed based on the Kimura two-parameter correction using ClustalX (Thompson et al., 1997), with gapped positions included.

3.4 Selection of strong transcriptional activators

A total of 22 positive co-transformations were identified. The relative strengths of activation of these positives were assessed using growth kinetics analysis and quantitative real-time PCR (QPCR). Assessing the growth rate of the positive clones in 3-AT selective media is an indirect

way of measuring transcription of the *HIS3* gene. Higher growth rates will correspond to an increased ability to synthesize histidine and therefore an increased ability to grow on selective media. The BacterioMatch II positive co-transformation described above provided a positive control. A pBT-MS2 and pTRG-var co-transformation that was picked from a non-selective screen plate and did not grow after incubation on a 2 mM selective screen plate was used as the negative control in these analyses. Glycerol stocks of each transcriptional activator and control co-transformation were streaked onto non-selective screen plates. These plates were incubated at 37^{o}C for 24 h. A single colony from each co-transformation was used to inoculate 2 ml of non-selective screen media and the inoculated cultures were incubated overnight at 37^{o}C with shaking at 215 rpm. For each sample, 10 μl of overnight culture was added to 150 μl of 2 mM 3-AT selective screen media in the wells of a 96-well plate. Each sample was prepared in triplicate. The plate was incubated at 37^{o}C with shaking and the absorbance at 600 nm of each well was read every 20 minutes for 8 h using a programmable plate reader (Synergy 2; BioTek, Winooski, VT). Growth curves of log10 (absorbance at 600 nm) against time were produced and mean growth rate constant, k, was calculated during logarithmic growth of each sample using the following equation:

Mean growth rate constant, $k = [\log10\ (A_{600})t_1 - \log10\ (A_{600})t_0] / 0.301t$

QPCR allows direct measurement of *HIS3* transcription relative to the amount of RNA transcriptional activator present. The negative control co-transformation was the same as described for the growth curve analysis. No positive control co-transformation was available since the normalizing RNA is the RNA transcriptional activator that was inserted into the target plasmid, which is absent from the BacterioMatch II kit positive control co-transformations. Glycerol stocks of each transcriptional activator and control co-transformation were streaked

onto non-selective screen plates. These plates were incubated at 37°C for 24 h. A single colony from each co-transformation was used to inoculate 2 ml of non-selective screen media and the inoculated cultures were incubated overnight at 37°C with shaking at 215 rpm. For each positive and control, 50 μl of overnight culture was added to 2 ml of 1 mM 3-AT selective screen media and incubated at 37°C with shaking at 215 rpm for 3 h. RNA was extracted from each culture as described above. RT-PCR reactions with reverse transcriptase omitted were used to confirm that there was no DNA contamination of the RNA isolations. Quantity and quality of RNA was determined by absorbance (Nanodrop spectrophotometer; Thermo Scientific, Wilmington, DE) and gel electrophoresis, respectively. One microlitre of total RNA was made into cDNA using the iScript select cDNA synthesis kit (Bio-Rad, Hercules, CA) following the manufacturer's protocol for random hexamer primed reactions. QPCR was performed using an ABI 7500 Fast real-time PCR machine in association with primers and probes designed using the Custom TaqMan gene expression assay service utilizing the TaqMan chemistry real-time PCR protocol (Applied Biosystems, Foster City, CA). Primers and FAM-labeled probes were designed to specifically recognize the cDNA sequences of *HIS3* (HIS3F, TGCTCTCGGTCAAGCTTTTAAAGA; HIS3R, CGCAAATCCTGATCCAAACCTTTT; HIS3FAM, CACGCACGGCCCCTAG) and the RNA transcriptional regulator (varRNAF, GCGGCTGGGAACGAAAC; varRNAR, CCACTAGTTCTAGCCGGAATTCTG; varRNAFAM, CCAATCGCAGCTCCCA). Briefly, QPCR reactions were set up using 10 μl 2x TaqMan Fast universal PCR mastermix (Applied Biosystems, Foster City, CA), 1 μl 20x TaqMan gene expression assay mix containing the primers and probe for either *HIS3* or the RNA transcriptional activator, 1 μl of cDNA, and 8 μl nuclease-free water per sample. Ten-fold serial dilutions of a representative cDNA sample were also prepared to assess the efficiency of the

reaction for each gene of interest. Samples were prepared in triplicate and loaded into a 96-well plate. Reactions were run using the following thermal profile: 95^{o}C for 20s; 40 cycles of 95^{o}C for 3s, 60^{o}C for 30s. Wells were analyzed during the anneal/extension step. Samples were analyzed using the ABI 7500 Fast system software utilizing the delta delta Ct method of relative quantification (Livak and Schmittgen, 2001). The relative fold difference of *HIS3* expression of each sample compared to the negative was normalized using RNA transcriptional activator expression. This method assumes optimum efficiency of each reaction. Reaction efficiency was calculated using the equation:

$$\text{Efficiency} = 10^{(-1/\text{slope})}$$

where slope is the slope of a graph of Ct value against log10(dilution). An efficiency of 2 is considered optimal since it indicates that the amplicon concentration is doubling with every amplification cycle. Average reaction efficiency was 2.020 (SE = 0.018). Correlation of the results of the growth curve analysis and the QPCR analysis was tested using a Spearman's rank correlation coefficient test. The Spearman's rank coefficienct, ρ, varies from 1 (perfect agreement between two rank orders) through 0 (rankings are completely independent) to -1 (one ranking is the reverse of the other). Confidence levels can be applied to ρ using published ρ statistical tables. Growth curve analysis and QPCR analysis gave statistically similar results (Spearmans Rank Correlation coefficient, ρ, of 0.689, $P<0.0005$).

RNA was extracted from representative samples deemed to be strong, medium, or weak transcriptional activators. Two RNA extractions were performed for each sample, each grown from a separate cfu from a streak of the glycerol stock of that sample. RNA extractions and QPCR analyses were performed as described above. Average reaction efficiency was 2.022 (SE = 0.014). In addition to comparing within each biological replicate, each individual Ct for HIS

and VAR of the three biological replicates of each sample were compared to the mean of the NEG Ct to calculate the delta delta Ct value. Thus there is a fold difference value for each combination of HIS Ct (n = 9) to VAR Ct (n = 9), therefore n = 81 for each RNA transcriptional activator. The means of the pooled fold difference values for each RNA transcriptional activator were compared by a one-way analysis of variance followed by Tukey's pairwise comparison (95% confidence interval). Individual pairs of data were compared by t test.

To verify that the four strong RNA transcriptional activators determined in the above experiments were true positives, their growth was determined on stringent 3-AT selective screen plates. Glycerol stocks of each positive and control co-transformation were used to inoculate 1.5 ml tubes containing 750 µl of non-selective screen media. The inoculated cultures were incubated at 37^{o}C for 30 minutes. One hundred microlitres of each culture was plated onto either 5 mM 3-AT selective screen plates, 3 mM 3-AT selective screen plates, or non-selective screen plates. These plates were incubated at 37^{o}C. Growth was assessed by visualizing the plates after 18 h, 24 h, and 32 h. In order to identify the strongest RNA transcriptional activator, their order in the growth kinetics assay and the compiled real-time PCR data was converted to a rank (with highest mean fold difference assigned to a rank of 1), as was their order in the selective screen assay. For the latter, rank was attributed to the rapidity of growth on the 5 mM 3-AT selective screen, with the fastest growth assigned a rank of 1.

3.5 Identification of interacting proteins by EMSA

Stabs from glycerol stocks of each sample and control co-transformation were used to inoculate 2 ml of non-selective screen media. However, the untransformed BacterioMatch II screening reporter strain competent cells were grown in M9+ HIS-dropout broth alone, and the pBT-MS2 alone transformation was grown in M9+ HIS-dropout broth + 25 µg/ml chloramphenicol. The

inoculated cultures were incubated overnight at 37^{o}C with shaking at 215 rpm. For each sample and control, 1 ml of overnight culture was added to 100 ml of their respective growth media and incubated overnight at 37^{o}C with shaking at 215 rpm. Total soluble proteins were extracted as described above. Soluble protein quantities were assessed using the BCA protein assay (Thermo Fisher Scientific Inc., Rockford, IL) and the Nanodrop spectrophotometer.

EMSAs were performed in triplicate using the LightShift Chemiluminescent EMSA kit (Pierce Themo Fisher Scientific Inc., Rockford, IL) following the manufacturer's protocol. A biotinylated RNA probe (5'BioVar266-288, Biotin-GAAGUUGGAUAUGGG) was synthesized (IDT DNA, Coralville, IA) that would complement the conserved 3' end of the RNA transcriptional activator at nucleotides 266 - 288. Reaction conditions were optimized using the LightShift EMSA optimization and control kit and RNase inhibitor was added to each binding reaction. The optimized binding conditions for probe 5'BioVar266-288 were as follows: 1X Binding Buffer; 50 ng/ml Poly (dI dC), 1 U/ml Superase-In (Ambion), 50 nM biotinylated probe, 2.5% glycerol, 5 mM $MgCl_2$, 0.05% NP-40, 10 mM EDTA, 100 μg soluble protein extraction. These conditions were used in all proceeding experiments. To investigate the specificity of binding, binding reactions were performed using the optimized probe conditions described above with and without a 100-fold excess of unbiotinylated probe. Native gel electrophoresis was performed using the NativePAGE Novex Bis-Tris gel system (Invitrogen, Carlsbad, CA). Duplicate gels were run. One of the gels was Coomassie stained. The duplicate native gel was transferred onto a nylon membrane, protein was cross-linked to the membrane by UV irradiation, and the probe was detected using the Chemiluminescent Nucleic Acid Detection Module. Bands of interest were matched to identical regions of the Coomasse stained gel, which were excised

and sequenced by mass spectrometry analysis (Cancer and Cell Biology Proteomics Core Facility, University of Cincinnati).

3.6 Immunoprecipitation of RNA transcriptional activator using RNA polymerase subunits

Immunoprecipitaion experiments were devised to investigate if the RNA transcriptional activator could be isolated using antibodies to rpoB and rpoD, two RNA polymerase subunits protein identified in the EMSA sequencing. The immunoprecipitation and analysis was carried using the Classic IP kit (Pierce Biotechnology, Rockford, IL) following the manufacturer's protocol, using the low-pH elution method to avoid denaturing conditions associated with the sample-buffer elution method. Briefly, the antibody is incubated with cell lysate to form an immune complex. The immune complex is captured using Protein A/G linked agarose beads in a minicolumn. After washing, the antibody-immune complex is eluted. The first wash and the elution was run on a SDS-PAGE gel and visualized using the SilverQuest staining kit (Invitrogen, Carlsbad, CA), and also used as template for PCR and reverse-transcriptase PCR using RTA-specific primers as described above. The control target plasmid, pTRG-control, was constructed as described for pTRG-var, except that a 10 nucleotide region that complemented the sequence opposite to the base of the stem-loop formed by the variable region in pTRG-var (TCTAGAGTCG), replaced the 40 nucleotide variable region. *E. coli* co-transformed with pBT-MS2 and pTRG-control were not able to grow on 1 mM 3-AT. To facilitate identification of immunoprecipitated RNA polymerase subunits, cell lysate from a co-RTA-3 culture was also spiked with 5 U of *E. coli* RNA polymerase holoenzyme (Epicentre, Madison, WI) before performing the immunoprecipitation.

3.7 Mutagenesis of the strongest transcriptional activator

The RTA-3 clone was mutated using the JBS dNTP-mutagenesis kit (Jena Bioscience, Jena, Germany) following the manufacturer's protocol. Briefly, RTA-3 was amplified from pTRG-RTA-3 in the presence of dNTPs and mutagenic dNTP analogs 8-oxo-dGTP and dPTP using primers RTRNAF and RanMut R (GCAGGCATGCGCGGCCGC, Tm 69.4^{o}C). Thirty PCR cycles were run to maximize the rate of mutagenesis. The mutagenic dNTP analogs are eliminated from the resultant amplicon by subjecting it to a second PCR containing natural dNTPs. Cloning, screening for transcriptional activation, and sequence identification of the mutants were performed as described above.

QPCR was used to determine the strength of transcriptional activation of the RTA-3 mutants. Briefly, RNA was extracted from clone RTA-3 and the RTA-3 mutant transcriptional activators. Three RNA extractions were performed for each sample, each grown from a separate cfu from a streak of the glycerol stock of that sample. RNA extractions and QPCR analyses were performed as described above. Average reaction efficiency was 2.015 (SE = 0.014). Each individual Ct for HIS and VAR of the three biological replicates of each sample were compared to the mean of the RTA-3 Ct to calculate the delta delta Ct value. Thus there is a fold difference value for each combination of HIS Ct (n = 9) to VAR Ct (n = 9), therefore n = 81 for each RNA transcriptional activator. The means of the pooled fold difference values for each RNA transcriptional activator were compared by a one-way analysis of variance followed by Tukey's pairwise comparison (95% confidence interval). Individual pairs of data were compared by t test.

Supporting Information

Supporting figures and tables are available at the bottom of this report.

REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. 215, 403-410.

Bernstein, D.S., Buter, N., Stumpf, C., and Wickens, M. (2002). Analyzing mRNA-protein complexes using a yeast three-hybrid system. Methods 26, 123-141.

Busby, S. and Ebright, R.H. (1994). Promoter structure, promoter recognition, and transcription activation in prokaryotes. Cell 79, 743-746

Bushman, F.D., Shang, C., and Ptashne, M. (1989). One glutamic acid residue plays a key role in the activation function of lambda repressor. Cell 58, 1163-1171

Buskirk, A.R., Kehayova, P.D., Landrigan, A., and Liu, D.R. (2003). In vivo evolution of an RNA-based transcriptional activator. Chem. Biol. 10, 533-540.

Buskirk, A.R., Landrigan, A., and Liu, D.R. (2004). Engineering a ligand-dependent RNA transcriptional activator. Chem. Biol. 11, 1157-1163.

Culler, S.J., Hoff, K.G., and Smolke, C.D. (2010). Reprogramming cellular behavior with RNA controllers responsive to endogenous proteins. Science 330, 1251-1255.

Harbaugh, S.V., Davidson, M.E., Chushak, Y.G., Kelley-Loughnane, N., and Stone, M.O. (2008). Riboswitch-based sensor in low optical background. Proc. of SPIE 7040, 70400C

Hunsicker, A., Steber, M., Mayer, G., Meitert, J., Klotzsche, M., Blind, M., Hillen, W., Berens, C., and Suess, B. (2009). An RNA aptamer that induces transcription. Chem. Biol. 16, 173-180.

Isaacs, F.J., Dwyer, D.J., Ding, C., Pervouchine, D.D., Cantor, C.R., and Collins, J.J. (2004). Engineered riboregulators enable post-transcriptional control of gene expression. Nat. Biotechnol. 22, 841-847.

Liu, C.C. and Arkin, A.P. (2010). The case for RNA. Science 330, 1185-1186

Livak, K.J. and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{(-\Delta\Delta Ct)}$ method. Methods 25, 402-408.

Lucks, B.J., Qi, L., Mutalik, V.K., Wang, D., and Arkin, A.P. (2011). Versatile RNA-sensing transcriptional regulators for engineering genetic networks. Proc. Nat. Acad. Sci. 108, 8617-8622.

Mattick, J.S. (2004). RNA regulation: a new genetics? Nat. Rev. Genet. 5, 316-323.

Novick R.P., Iordanescu, S., Projan, S.J., Kornblum, J., and Edelman, I (1989). pT181 plasmid replication is regulated by a coutertranscript-driven transcriptional attenuator. Cell 59, 395-404.

Ptashne, M. and Gann, A. (1997). Transcriptional activation by recruitment. Nature 386, 569-577.

Saha, S., Ansari, A.Z., Jarell, K.A., and Ptashne, M. (2003). RNA sequences that work as transcriptional activating regions. Nucleic Acids Res. 31, 1565-1570.

Storz, G. (2002). An expanding universe of noncoding RNAs. Science 296, 1260-1263.

Storz, G., Altuvia, S., and Wassarman, K.M. (2005). An abundance of RNA regulators. Annu. Rev. Biochem. 74, 199-217.

Storz, G., Opdyke, J.A., and Wassarman, K.M. (2006). Regulating bacterial transcription with small RNAs. Cold Spring Harb. Symp. Quant. Biol. 71, 269-273.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876-4882.

Topp, S. and Gallivan, J.P. (2010). Emerging applications of riboswitches in chemical biology. ACS Chem. Biol. 5, 139-148.

Wang, S., Shepard, J.R.E., and Shi, H. (2010). An RNA-based transcription activator derived from an inhibitory aptamer. Nucleic Acids Res. 38, 2378-2386.

Wassarman, K.M. and Storz, G. (2000). 6S RNA regulates E. coli RNA polymerase activity. Cell 101, 613-623.

Waters, L.S. and Storz, G. (2009). Regulatory RNAs in bacteria. Cell 136, 615-628.

Windbichler, N., von Pelchrzim, F., Mayer, O., Csaszar, E., and Schroeder, R. (2008). Isolation of small RNA-binding proteins from *E. coli*. RNA Biol. 5, 1-11.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 31, 3406-3415.

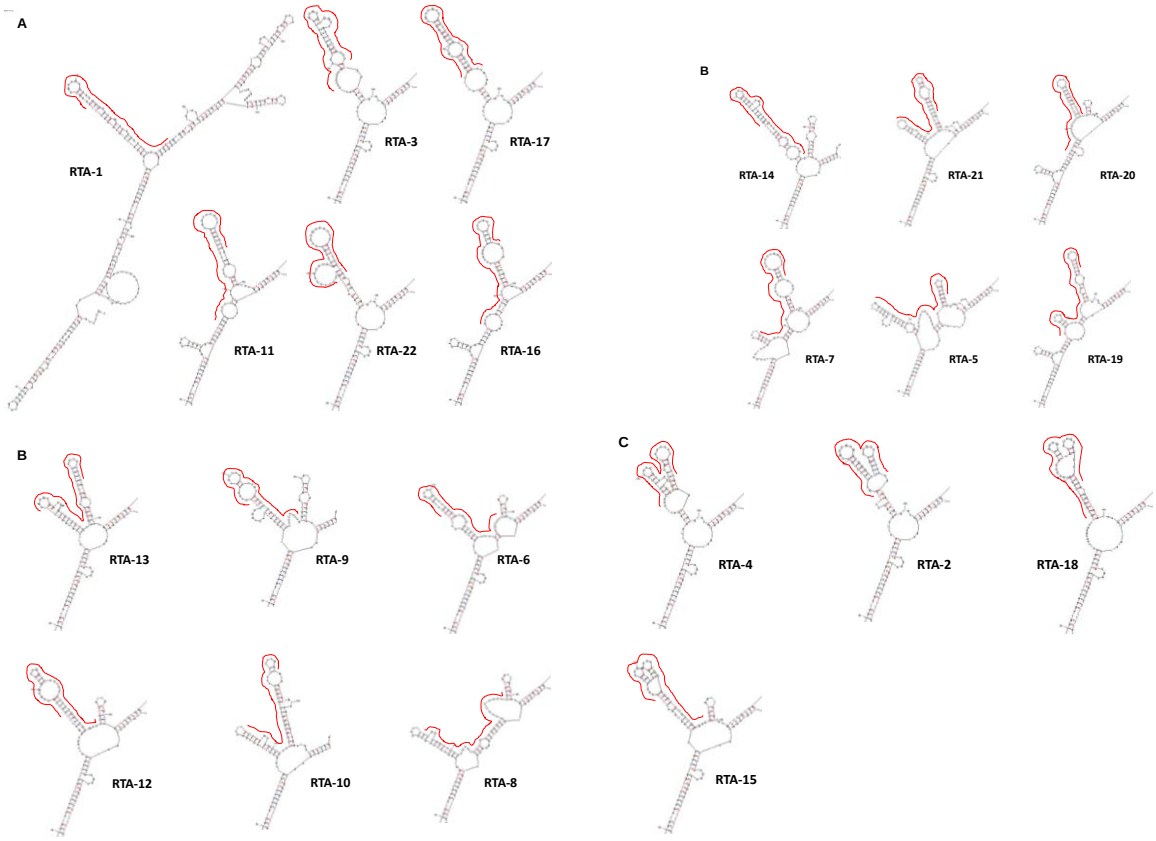


Figure S1. Predicted secondary structure of the positive transcriptional activators. The regions exhibiting conserved secondary structure are omitted in all except RTA-1. A) Variable region possessing a single stem-loop B) Variable region possessing a double stem-loop. C) Variable region possessing paired stem-loops. Variable regions indicated by red outline and capital letters. Prediction made using mfold.

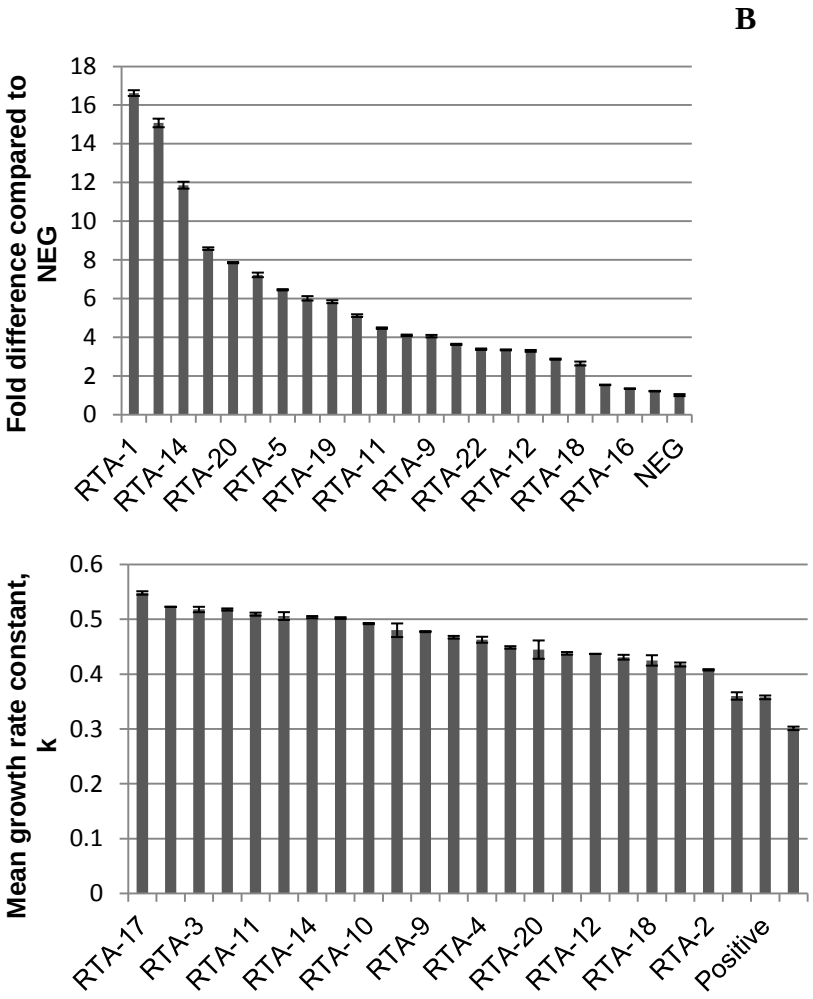


Fig. S2 Assessment of the strength of transcriptional activation. A) Average fold difference of *HIS3* expression compared to a clone that did not grow on 2 mM 3-AT selective screen plates (negative; NEG). *HIS3* expression was normalized to RNA transcriptional activator expression. Fold difference was compared using the delta delta Ct method. Each individual Ct value of three replicates per clone was compared to the mean NEG Ct to calculate each delta delta Ct value (n=9). Data are represented as mean +/- SEM. B) Average growth rate constant, k, of triplicated samples of positive transcriptional activators. Data are represented as mean +/- SEM.

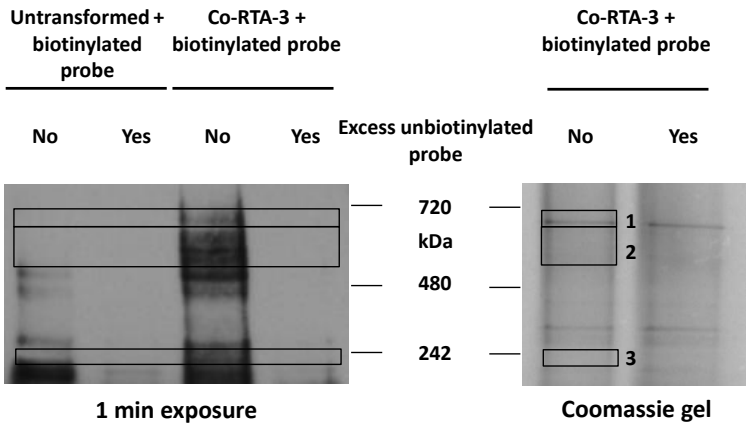


Fig. S3. **EMSA to identify proteins interacting with the RNA transcriptional activator.** A) A biotinylated probe specific to the RNA transcriptional activator was crosslinked to the activator and a shift in electrophoretic mobility was observed using native gel electrophoresis. Untransformed bacteria were used to indicate non-specific probe binding. Excess unbiotinylated probe was used to verify probe specificity. Bands only present in the co-transformed bacteria, corresponding to regions 1-3, were sequenced. Gel representative of three independent experiments.

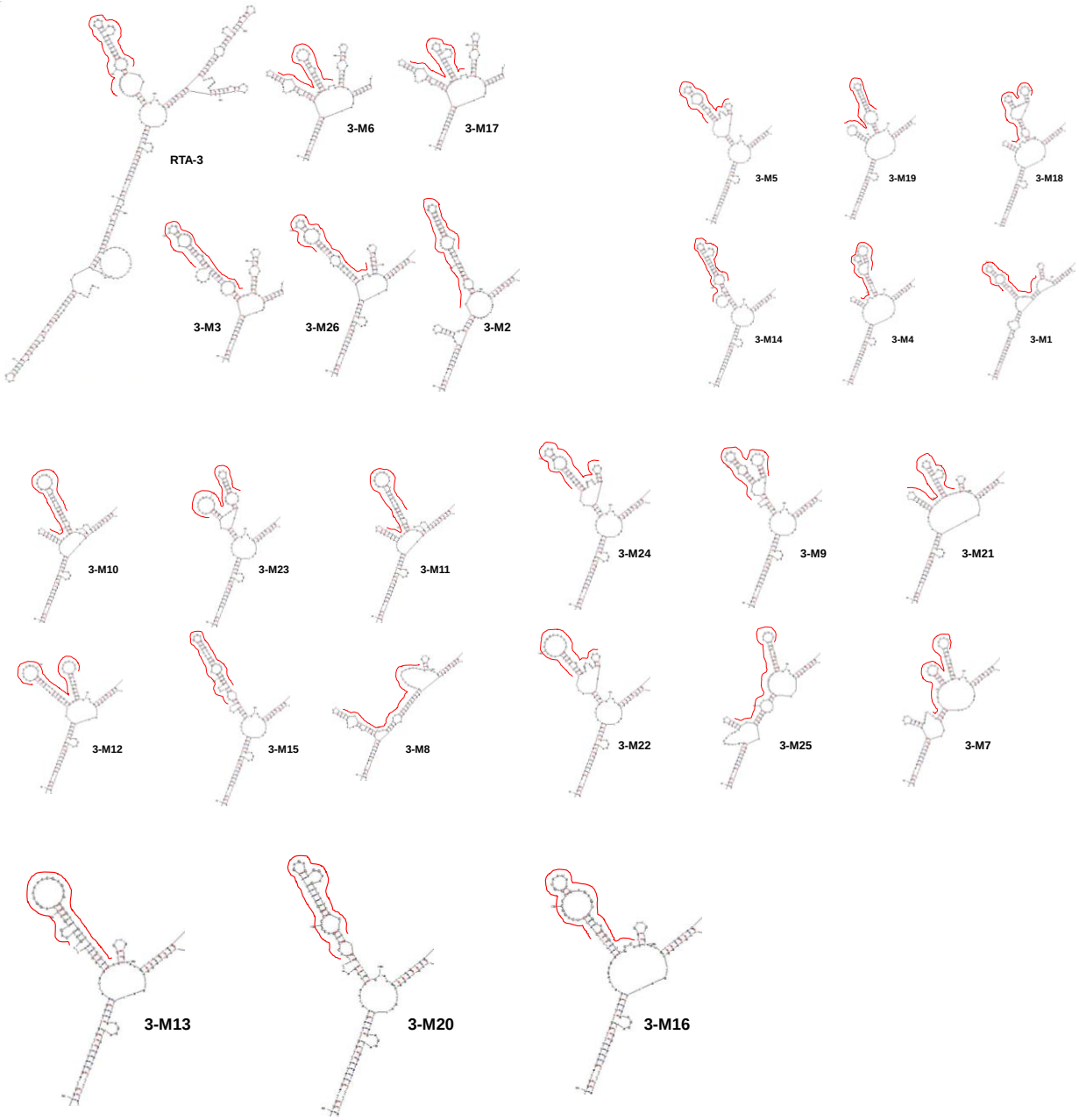


Figure S4. Predicted secondary structure of the RTA-3 mutant activators. To facilitate comparison, RTA-3 is shown in its entirety while the regions exhibiting conserved secondary structure are omitted from the RTA-3 mutants. RTA-3 mutants are listed in the same order as Frig. 5. Variable regions indicated by red outline and capital letters. Prediction made using mfold.

Group by stem loop as in Fig. S1, group in numerical or strength order within stem-loop group?

Table S1. Proteins identified by sequencing regions 1, 2, and 3 of the EMSA gel. Members of the RNA polymerase holoenzyme and the lambda repressor are highlighted.

EMSA protein hits		Region 1		Gene ontology from UniProt[2]
Hit #	Accession #	Protein	Protein score[1]	Biological process
1	gi|45686198	groEL	1734	protein refolding
2	gi|409789	RNA polymerase beta (rpoB)	1015	transcription
3	gi|145544	Colicin I receptor	860	Ion Transport
4	gi|1311039	Chain A, Dipeptide Binding Protein Complex With Glycyl-L-Leucine	827	Peptide Transport
5	gi|6730010	Chain A, Ferric Enterobactin Receptor	506	ion transport
6	gi|15799879	DL-methionine transporter subunit	472	AA transport
7	gi|110807172	elongation factor Tu	457	Protein biosynthesis/antibiotic resistance
8	gi|16131416	Dipeptide transporter	420	transport
9	gi|15988003	Chain A, Crystal Structure Of Outer Membrane Protease Ompt From Escherichia Coli	359	Proteolysis
10	gi|38704050	fructose-bisphosphate aldolase	357	Glycolysis
11	gi|193062142	outer membrane protein TolC	344	protein transport
12	gi|16130963	RNA polymerase, sigma 70 (sigma D) factor	341	transcription
13	gi|284822073	outer membrane receptor FepA	302	ion transport
14	gi|16130232	phosphate acetyltransferase	278	protein binding acetlytransferase
15	gi|15801691	putative receptor	252	receptor
16	gi|170682536	ATP-dependent metallopeptidase HflB	244	proteolysis
17	gi|15803822	DNA-directed RNA polymerase subunit alpha	242	transcription
18	gi|15802857	lysine-, arginine-, ornithine-binding periplasmic protein	240	transport
19	gi|16130152	outer membrane porin protein C	194	ion transport
20	gi|51235578	outer membrane protein A	187	ion transport
21	gi|187776372	acriflavine resistance protein A	176	protein transport
22	gi|16129867	cystine transporter subunit	150	transport

23	gi|15800457	peptidoglycan-associated outer membrane lipoprotein	143	protein binding
24	gi|19548975	aspartokinase I-homoserine dehydrogenase I	138	AA biosynthesis
25	gi|37953544	isocitrate dehydrogenase	134	TCA cycle
26	gi|223138	protein I,membrane	132	transport
27	gi|15800456	translocation protein TolB	125	Protein transport
28	gi|188493649	outer membrane assembly lipoprotein YfgL	102	protein binding
29	gi|15804741	entericidin B membrane lipoprotein	82	response to toxin
30	gi|256025393	Acetolactate synthase 3 regulatory subunit	75	AA biosynthesis
31	gi|15802089	Murein lipoprotein	74	protein binding/lipid modification
32	gi|16129736	scaffolding protein for murein synthesizing machinery	74	scaffold
33	gi|157149414	TraT complement resistance protein	74	Pathogenesis/Virulence
34	gi|15803639	hypothetical protein Z4452	73	Unknown
35	gi|15802150	OsmE	71	cell membrane/DNA-binding/transcription
36	gi|15803767	cytochrome d ubiquinol oxidase subunit III	69	electron transport
37	gi|26250246	adhesin	66	Pathogenesis/Virulence

[1]Protein score is a function of the number of peptides that matched the same protein in the NCBI database. Only those proteins with significant matches (P<0.005) are shown.

[2]Jain, E., Bairoch, A., Duvaud, S., Phan, I., Redaschi, N., Suzek, B.E., Martin, M.J., McGarvey, P., and Gasteiger, E. (2009). Infrastructure for the life sciences: design and implementation of the UniProt website. BMC Bioinformatics 10, 136